

## University of Groningen

### Multicenter CT phantoms public dataset for radiomics reproducibility tests

Kalendralis, Petros; Traverso, Alberto; Shi, Zhenwei; Zhovannik, Ivan; Monshouwer, Rene; Starmans, Martijn P. A.; Klein, Stefan; Pfaehler, Elisabeth; Boellaard, Ronald; Dekker, Andre

*Published in:*  
Medical Physics

*DOI:*  
[10.1002/mp.13385](https://doi.org/10.1002/mp.13385)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2019

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Kalendralis, P., Traverso, A., Shi, Z., Zhovannik, I., Monshouwer, R., Starmans, M. P. A., Klein, S., Pfaehler, E., Boellaard, R., Dekker, A., & Wee, L. (2019). Multicenter CT phantoms public dataset for radiomics reproducibility tests. *Medical Physics*, 46(3), 1512-1518. <https://doi.org/10.1002/mp.13385>

#### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

#### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

## DAILY QA

*Finished before your first cup of coffee*

## MONTHLY QA

*Never re-learn workflow again*

## ANNUAL QA

*Confidence with fewer gray hairs*

## STEREOTACTIC

*When millimeters matter most*

## ABSOLUTE DOSIMETRY

*"Pinpoint" is our margin of error*

## PATIENT DOSIMETRY

*Pre-treatment is just the beginning*



*QA Solutions.*

*Our expertise is at your service.*

**STANDARDIMAGING®**



MORE THAN JUST QA PRODUCTS, FIND QA SOLUTIONS AT

**[www.standardimaging.com](http://www.standardimaging.com)**

# Multicenter CT phantoms public dataset for radiomics reproducibility tests

Petros Kalendralis,<sup>a)</sup> Alberto Traverso, and Zhenwei Shi

*MAASTRO Clinic and School for Oncology and Development Biology (GROW), Maastricht University Medical Centre+, 6229 ET Maastricht, The Netherlands*

Ivan Zhovannik

*MAASTRO Clinic and School for Oncology and Development Biology (GROW), Maastricht University Medical Centre+, 6229 ET Maastricht, The Netherlands*

*Department of Radiation Oncology, Radboud University Medical Center, 6525 GC Nijmegen, The Netherlands*

René Monshouwer

*Department of Radiation Oncology, Radboud University Medical Center, 6525 GC Nijmegen, The Netherlands*

Martijn P. A. Starmans and Stefan Klein

*Department of Radiology and Nuclear Medicine, Erasmus Medical Centre, 3015 GD Rotterdam, The Netherlands*

*Department of Medical Informatics, Erasmus Medical Centre, 3015 GD Rotterdam, The Netherlands*

Elisabeth Pfaehler

*University Medical Center Groningen, 9713 GZ Groningen, The Netherlands*

Ronald Boellaard

*Department of Radiology and Nuclear Medicine, VU Medical Center, 1081 HV Amsterdam, The Netherlands*

Andre Dekker, and Leonard Wee

*MAASTRO Clinic and School for Oncology and Development Biology (GROW), Maastricht University Medical Centre+, 6229 ET Maastricht, The Netherlands*

(Received 21 August 2018; revised 15 November 2018; accepted for publication 6 December 2018; published 29 January 2019)

**Purpose:** The aim of this paper is to describe a public, open-access, computed tomography (CT) phantom image set acquired at three centers and collected especially for radiomics reproducibility research. The dataset is useful to test radiomic features reproducibility with respect to various parameters, such as acquisition settings, scanners, and reconstruction algorithms.

**Acquisition and validation methods:** Three phantoms were scanned in three independent institutions. Images of the following phantoms were acquired: Catphan 700 and COPDGene Phantom II (Phantom Laboratory, Greenwich, NY, USA), and the Triple modality 3D Abdominal Phantom (CIRS, Norfolk, VA, USA). Data were collected at three Dutch medical centers: MAASTRO Clinic (Maastricht, NL), Radboud University Medical Center (Nijmegen, NL), and University Medical Center Groningen (Groningen, NL) with scanners from two different manufacturers Siemens Healthcare and Philips Healthcare. The following acquisition parameter were varied in the phantom scans: slice thickness, reconstruction kernels, and tube current.

**Data format and usage notes:** We made the dataset publically available on the Dutch instance of “Extensible Neuroimaging Archive Toolkit-XNAT” (<https://xnat.bmia.nl>). The dataset is freely available and reusable with attribution (Creative Commons 3.0 license).

**Potential applications:** Our goal was to provide a findable, open-access, annotated, and reusable CT phantom dataset for radiomics reproducibility studies. Reproducibility testing and harmonization are fundamental requirements for wide generalizability of radiomics-based clinical prediction models. It is highly desirable to include only reproducible features into models, to be more assured of external validity across hitherto unseen contexts. In this view, phantom data from different centers represent a valuable source of information to exclude CT radiomic features that may already be unstable with respect to simplified structures and tightly controlled scan settings. The intended extension of our shared dataset is to include other modalities and phantoms with more realistic lesion simulations. © 2019 The Authors Medical Physics published by Wiley Periodicals, Inc. on behalf of American Association of Physicists in Medicine. [<https://doi.org/10.1002/mp.13385>]

## 1. INTRODUCTION

Computer-aided analysis of clinical radiological images offers a data-at-large-scale approach toward personalized medicine<sup>1</sup> wherein tumor phenotype may be inferred using images of the entire tumor instead of selective sample

biopsies. On the premise that phenotypic variability affects clinical outcome,<sup>2</sup> medical imaging offers an efficient and noninvasive method to determine prognosis.

This approach has immense potential to support clinical decision-making in the personalized medicine paradigm,<sup>3</sup> that is, which would be a superior choice of treatment for a

given person. Studies in the active field of image-derived markers (i.e., “radiomics”) strongly suggest that tomographic images do indeed embed more prognostic information than may be seen by an unassisted human eye.<sup>4–8</sup> In order to be widely generalizable and have meaningful clinical use, it is essential that reproducibility of features can be tested in phantoms,<sup>9,10</sup> in addition to validating models in human subjects across different settings and multiple independent institutions.<sup>11–13</sup>

Studies have shown that feature reproducibility may be affected by differences in image acquisition parameters, such as slice thickness and reconstruction algorithm.<sup>14–17</sup> Since clinical image acquisition protocols are one of the major sources of variation among different hospitals, phantoms allow testing, comparison, and harmonization of radiomic features in similar vein to diagnostic imaging quality assurance. We hypothesize that even simplified phantoms allow us to test for radiomic features that may already become unstable even under tightly constrained conditions.

In this data publication, we offer computed tomography (CT) scans of simple phantoms across three Dutch academic medical centers for open access. We chose to start with CT since this modality is readily available in many centers and is a workhorse imaging modality for radiotherapy intervention planning. In many clinics, CT scanners are mature technology with well-established protocols for calibration, quality assurance, and routine maintenance.

## 2. ACQUISITION AND VALIDATION METHODS

### 2.A. Phantoms

#### 2.A.1. Catphan 700

To obtain a baseline for overall CT scanner performance, we scanned a Catphan 700 phantom (Phantom Laboratory, Greenwich, NY, USA) that had been designed specifically for routine quality assurance on CT scanners. It is only suitable for use in CT, and contains test modules for contrast, geometric accuracy, and spatial resolution.<sup>18,19</sup>

#### 2.A.2. COPDGene Phantom II

The COPDGene Phantom II (Phantom Laboratory, Greenwich, NY, USA) was designed for thoracic CT quality assurance in prospective clinical trials (specifically asthma and chronic obstructive pulmonary disorder) with guidance from the Quantitative Image Biomarker Alliance Technical Committee. We used the CCT162 version, which included the standard version CTP698 with two additional supports and acrylic end-plates for stabilization of the phantom during the scanning. An outer polyurethane ring simulated tissue attenuation while an internal oval body (15 cm × 25 cm) simulated lung attenuation. The inner oval held a number of cylindrical cavities for foam, acrylic, and water,<sup>20,21</sup> as well as a number of internal structures simulating different-sized bronchi.

#### 2.A.3. Triple modality 3D Abdominal Phantom

A 3D multimodality Abdominal Phantom (CIRS, Norfolk, Virginia, USA) measuring 26 cm × 12.5 cm × 19 cm<sup>22</sup> was designed to be used for liver biopsy training under guidance by CT, magnetic resonance imaging, or ultrasonography. We scanned Model 057A that simulated the abdomen of a small adult. The materials encased within the phantom represented the liver, portal vein, kidneys, bottom of the lungs, abdominal aorta, vena cava, lumbar spine, and six lowest ribs.

### 2.B. Image acquisition

The images used in our study were acquired using three different CT scanners at independent Dutch centers: MAAS-TRO Clinic (Maastricht), Radboud University Medical Center (Nijmegen) and University Medical Center Groningen (Groningen). The standard clinical operating procedures for thoracic and abdominal radiotherapy planning CT scans at each of the three centers were used to generate a baseline scan of each phantom. These baseline parameters are stated in Tables I and II, for the Phantom Laboratory and CIRS phantoms, respectively.

We subsequently applied perturbations to imaging settings of the baseline scan. We adjusted the following parameters strictly one at a time and saved each scan: slice thickness (1, 3, and 5 mm), reconstruction kernels (between three and five settings depending on the scanner), and current-exposure product (50, 150, and 300 mAs). The individual setting for each scan is given in Tables III and IV, for the Phantom Laboratory and CIRS phantoms, respectively.

### 2.C. Image annotations

CatPhan 700 images were only used for image quality assessment of the baseline scans between participating centers, therefore, no annotations were added to the scans.

Regions of interest (ROIs) on the COPDGene and Abdominal Phantoms were manually delineated in MIRADA DBx (version 1.2.0.59, Mirada Medical, Oxford, United Kingdom). In the COPD phantom, we delineated four distinct spherical ROIs within two of the insert cavities. In the multimodality phantom, we delineated two different ROIs corresponding to two of the simulated liver lesions, one large and one small (as shown in Fig. 1). The delineations were performed by one medical physicist at MAASTRO Clinic. All images and annotations were then exported as Digital Imaging and Communications in Medicine (DICOM)-Radiotherapy (RT) objects.

### 2.D. Data format and usage notes

Our scans are made open access via an instance of the Extensible Neuroimaging Archive Toolkit (XNAT) hosted within Dutch national research infrastructure (TraIT,

TABLE I. CT scanner details and image acquisition parameters for baseline scans of the Catphan 700 and COPDGene Phantom II in each of the participating clinics.

Parameters	DICOM tags	MAASTRO Clinic (MAAS)	Radboud University Medical Center (RADB)	University Medical Center Groningen (UMCG)
Catphan 700/COPDGene Phantom II baseline scan parameters				
Manufacturer	(0008, 0070)	Siemens	Phillips	Siemens
Model	(0008, 1090)	Biograph 40	Brilliance Big Bore	Biograph 64
Software Version	(0018, 1020)	syngo CT 2006A	3.6.6	VG60A
Slice thickness (mm)	(0018, 0050)	3	3	3
TUBE VOLTAGE (KV)	(0018, 0060)	120	120	80
Reconstruction diameter (mm)	(0018, 1100)	500	255	239
Tube current (mA)	(0018, 1151)	39	134	149
EXPOSURE (mAs)	(0018, 1152)	24	124	53
Convolution kernel	(0018, 1210)	B31f	B	I30f
Rows	(0028, 0010)	512	1024	512
Columns	(0028, 0011)	512	1024	512
Pixel spacing	(0028, 0030)	0.98	0.25	0.46
Bits stored	(0028, 0101)	12	12	12
High bit	(0028, 0102)	11	11	11
Rescale offset	(0028, 1052)	−1024	−1024	−1024
Rescale slope	(0028, 1053)	1	1	1

TABLE II. CT scanner details and image acquisition parameters for baseline scans of the multimodality CIRS Abdominal Phantom in each of the participating clinics.

Parameters	DICOM tags	MAASTRO Clinic (MAAS)	Radboud University Medical Center (RADB)	University Medical Center Groningen (UMCG)
Triple modality 3D Abdominal Phantom baseline scan parameters				
Manufacturer	(0008, 0070)	Siemens	Phillips	Siemens
Model	(0008, 1090)	Biograph 40	Brilliance Big Bore	Biograph 64
Software Version	(0018, 1020)	syngo CT 2006A	3.6.6	VG60A
Manufacturer	(0008, 0070)	Siemens	Phillips	Siemens
TUBE VOLTAGE (KV)	(0018, 0060)	120	120	80
Reconstruction diameter (mm)	(0018, 1100)	500	255	239
Tube current (mA)	(0018, 1151)	118	190	18
EXPOSURE (mAs)	(0018, 1152)	73	175	9
Convolution kernel	(0018, 1210)	B30f	B	I30f
Rows	(0028, 0010)	512	512	512
Columns	(0028, 0011)	512	512	512
Pixel spacing	(0028, 0030)	0.98	0.75	0.59
Bits stored	(0028, 0101)	12	12	12
High bit	(0028, 0102)	11	11	11
Rescale offset	(0028, 1052)	−1024	−1024	−1024
Rescale slope	(0028, 1053)	1	1	1

www.ctmm-trait.nl).<sup>23</sup> XNAT is an open source platform for imaging-based research and clinical investigations, which manages access to different datasets compartmentalized into separate projects (i.e., collections). Within each collection, XNAT permits browsing of individual cases. The platform supports direct uploading of DICOM images and DICOM-RT objects (plan, structure set, and dose

grid) with *http* file transfer.<sup>24</sup> Studies in XNAT can be queried and retrieved by means of an API (Application Programming Interface) in the Python programming language by installing the *xnat* library (<https://pypi.org/project/xnat/>).

The Phantom Laboratory COPD phantom images have been uploaded to the XNAT collection STWSTRATEGY-



TABLE III. The individual scan settings for the Catphan 700 and COPD II phantoms from the participating different Dutch clinics.

Subject	Institution	Slice thickness (mm)	Voltage (kvp)	Current (mA)	Exposure (mAs)	Convolution kernel
Collection: series 1 — Catphan 700 and COPD II individual subject scan settings						
CatPhan-01-MAAS	MAASTRO	3	120	39	24	B31f
CatPhan-01-RADB	Radboud	3	120	134	124	B
CatPhan-01-UMCG	Groningen	3	80	165.5	58.5	I30f
COPD-001-MAAS	MAASTRO	3	120	130	80.5	B31f
COPD-001-RADB	Radboud	3	120	210	194	B
COPD-001-UMCG	Groningen	3	120	191	68	I30f
COPD-002-MAAS	MAASTRO	1	120	112.5	69.5	B31f
COPD-002-RADB	Radboud	1	120	210	194	B
COPD-002-UMCG	Groningen	1	120	205	73	I30f
COPD-003-MAAS	MAASTRO	5	120	106.5	66	B31f
COPD-003-RADB	Radboud	5	120	210	194	B
COPD-003-UMCG	Groningen	5	120	195	69	I30f
COPD-004-MAAS	MAASTRO	3	120	91	56	B31f
COPD-004-RADB	Radboud	3	120	54	50	B
COPD-004-UMCG	Groningen	3	120	140	50	I30f
COPD-005-MAAS	MAASTRO	3	120	80	50	B31f
COPD-005-RADB	Radboud	3	120	108	100	B
COPD-005-UMCG	Groningen	3	120	280	100	I30f
COPD-006-MAAS	MAASTRO	3	120	130	80.5	B41f
COPD-006-RADB	Radboud	3	120	325	300	B
COPD-006-UMCG	Groningen	3	120	660	300	I30f
COPD-007-MAAS	MAASTRO	3	120	130	80.5	B41f
COPD-007-RADB	Radboud	3	120	210	194	A
COPD-007-UMCG	Groningen	3	100	230	104	I40f
COPD-008-MAAS	MAASTRO	3	120	130	80.5	B75f
COPD-008-RADB	Radboud	3	120	210	194	C
COPD-008-UMCG	Groningen	3	100	231	104	I44f
COPD-009-MAAS	MAASTRO	3	120	130	80.5	B60f
COPD-009-RADB	Radboud	3	120	210	194	E
COPD-009-UMCG	Groningen	3	100	236	107	I49f
COPD-010-MAAS	MAASTRO	3	120	130	80.5	B80f
COPD-010-RADB	Radboud	3	120	210	194	L
COPD-010-UMCG	Groningen	3	100	232	105	I50f
COPD-011-UMCG	Groningen	3	100	238	108	I70f
COPD-012-UMCG	Groningen	3	100	236	107	B30f

Phantom\_Series1: (<https://xnat.bmia.nl/data/projects/stwstrategyps1>).

The CIRS multimodality Abdominal Phantom images have been uploaded to the XNAT collection STWSTRATEGY-Phantom\_Series2: (<https://xnat.bmia.nl/data/projects/stwstrategyps2>).

The Phantom Laboratories Catphan 700 phantom images have been uploaded to the XNAT collection STW-STRATEGY-Phantom\_Series3: (<https://xnat.bmia.nl/data/projects/stwstrategyps3>).

In each of the above collections, the subject identifier matches exactly the names shown in the leftmost column of Tables III and IV. DICOM-formatted images and the annotations as DICOM-RTStruct objects are nested under the subject level. A python script for downloading an entire

collection is available here: (<https://github.com/maastrolinix/XNAT-collections-download-script>).

The images of the Catphan 700 quality assurance phantom from each center were analyzed online on the quality assurance tests webpage of the ImageOwl company (<https://catphanqa.imageowl.com/>). Clinical lung CT imaging protocols were used as the reference baseline for radiomics studies, rather than the vendors' service scan setting. The ImageOWL vendor service provides a detailed quality assurance analysis along with the implementation of linearity and sensitometry plots, noise measurements, and the spatial linearity. All quality assurance test parameters were within tolerance for the clinical lung scan settings used as the reference. The quality assurance reports can be found in Data S1.

TABLE IV. The individual settings of the Triple modality 3D Abdominal Phantoms from the three participating Dutch clinics.

Subject	Institution	Slice thickness (mm)	Voltage (kvp)	Current (mA)	Exposure (mAs)	Convolution kernel
Collection: series 2 — CIRS multimodality phantom individual subject scan settings						
CIRS-AB-001-MAAS	MAASTRO	3	120	118	73	B30f
CIRS-AB-001-RADB	Radboud	3	120	190	175	B
CIRS-AB-001-UMCG	Groningen	3	100	100	50	I30f
CIRS-AB-002-MAAS	MAASTRO	1	120	133	83	B30f
CIRS-AB-002-RADB	Radboud	1	120	190	175	B
CIRS-AB-002-UMCG	Groningen	1	100	95	47	I30f
CIRS-AB-003-MAAS	MAASTRO	5	120	136	85	B30f
CIRS-AB-003-RADB	Radboud	5	120	190	175	B
CIRS-AB-003-UMCG	Groningen	5	100	98	49	I30f
CIRS-AB-004a-UMCG	Groningen	3	120	100	50	I30f
CIRS-AB-004b-UMCG	Groningen	3	120	100	50	I30f
CIRS-AB-004-MAAS	MAASTRO	3	120	141	88	B30f
CIRS-AB-004-RADB	Radboud	3	120	54	50	B
CIRS-AB-005a-UMCG	Groningen	3	120	200	100	I30f
CIRS-AB-005b-UMCG	Groningen	3	120	200	100	I30f
CIRS-AB-005-MAAS	MAASTRO	1	120	137	85	B30f
CIRS-AB-005-RADB	Radboud	3	120	108	100	B
CIRS-AB-006-MAAS	MAASTRO	5	120	137.5	85.5	B30f
CIRS-AB-006-RADB	Radboud	3	120	325	300	B
CIRS-AB-006-UMCG	Groningen	3	120	600	300	I30f
CIRS-AB-007-RADB	Radboud	3	120	190	175	A
CIRS-AB-007-UMCG	Groningen	3	100	98	49	I40f
CIRS-AB-008-RADB	Radboud	3	120	190	175	C
CIRS-AB-008-UMCG	Groningen	3	100	98	49	I44f
CIRS-AB-009-RADB	Radboud	3	120	190	175	D
CIRS-AB-009-UMCG	Groningen	3	100	96	48	I49f
CIRS-AB-010-UMCG	Groningen	3	100	97	48	I50f
CIRS-AB-011-UMCG	Groningen	3	100	98	49	I70f
CIRS-AB-012-UMCG	Groningen	3	100	97	48	B30f

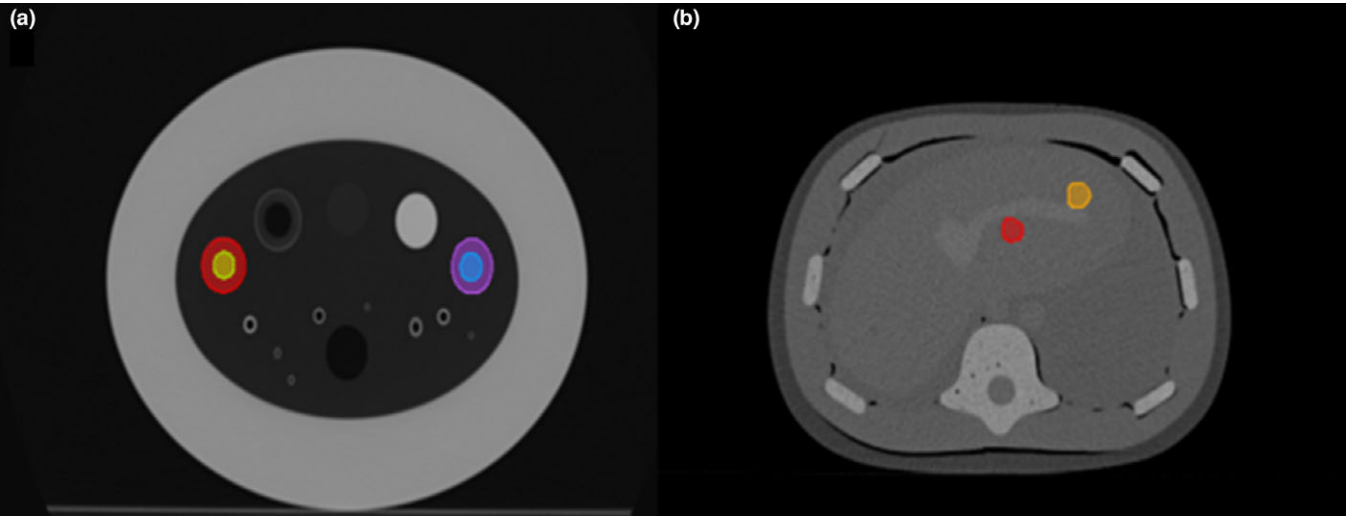


FIG. 1. The delineated spherical ROIs within two of the inserts cavities for the COPD and Triple modality 3D Abdominal Phantoms are presented in (a) and (b), respectively.

### 3. DISCUSSION

We have made publically available multicenter phantom CT scans to support investigations in radiomics repeatability and reproducibility, specifically to identify features that may be unstable with respect to image acquisition settings in simplified geometry.

Radiomics reproducibility may be investigated as a function of: scanner manufacturer/scanner type, slice thickness, tube current (i.e., signal to noise ratio), and reconstruction algorithms. We invite the radiomics community to make use of our dataset for research by extracting radiomic features with their own processing pipelines and comparing the results with other investigators. We also invite the community to contact us in order to share the results of their computations. For the next steps, we intend to host the computed features set from the open source library *pyradiomics* v2 (<https://github.com/Radiomics/pyradiomics>)<sup>25</sup> as well as the associated DICOM image metadata on a public open-access website ([www.radiomics.org](http://www.radiomics.org)).

This is a fundamental step toward improving benchmarking and standardization of the radiomics field of study. This is in support of valuable harmonization projects such as the IBSI (Image Biomarker Standardization Initiative).<sup>26</sup> The features and metadata will be made available as linked Resource Descriptor Format (RDF) objects labeled with a dedicated radiomic-specific semantic web ontology (<https://bioportal.bioontology.org/ontologies/RO>), such that the data can be queried through the SPARQL language. To assist the radiomics community with data sharing, a standard tabular template and conversion script to RDF will also be provided at [www.radiomics.org](http://www.radiomics.org).

A number of key limitations in the data must be noted at the present time. First, as explicitly declared by the phantom manufacturers, the phantoms used in this study had not been designed with the specific aim of simulating standard radiomic features. It is presently not fully understood exactly what should be used as a canonical set of imaging features.

Secondly, we posit that the so-called “test lesions” within the current phantoms represent oversimplified geometries and relatively uniformly dense material. Complex texture patterns and shape features are not well represented in such simple phantoms. However, these phantoms do present a preliminary opportunity for investigating reproducibility of radiomic features, thus we may be able to test for certain features that already unstable in simplified conditions. We would assert that a feature that is not reproducible in such a constrained setting might be unlikely to be highly reproducible in multi-institutional human studies. To improve on the current situation, the dataset might be expanded by scans of more phantoms that contain more realistic tumor-mimicking inserts. These may prove to be more suitable for selecting stable features for inclusion in radiomic investigations.

One example of a public phantom dataset which is available on “The Cancer Imaging Archive-TCIA” is the Credence Cartridge Radiomics (CCR) Phantom (<https://wiki.cancerimagingarchive.net/display/Public/Credence+Cartridge+Rad>

omics+Phantom+CT+Scans). The CCR phantom collection has a similar goal as our study, the investigation of the reproducibility of radiomic features. There is a significant factor that differentiates the CCR phantom public dataset from our phantoms public collections. The structure of the CCR phantom which includes ten cartridges, each with a unique texture, addresses only the question of repeatability and reproducibility of textural features.

Lastly, while we have started with CT as the most commonly available imaging modality in our field, we intend to expand this collection to include positron emission tomography (PET) and magnetic resonance imaging (MRI).

In addition to making available multicenter and multi-modality phantoms for radiomics reproducibility studies, future work in this field should make publicly accessible DICOM metadata and image preprocessing steps, so as to make radiomics studies as findable, accessible, interoperable, reusable (FAIR) as possible. To this end, image metadata needs to be linked to the features using publicly available Semantic DICOM (SEDI) ontology<sup>27</sup> and the Radiomics ontology needs to be extended to cover image preprocessing.

### 4. CONCLUSION

We offer a publicly accessible multicenter CT phantom dataset with carefully controlled image acquisition parameters to support reproducibility research in the field of radiomics. The dataset is hosted in a well-established and publicly funded XNAT instance. The data are shared under a Creative Commons Attribution 3.0 License (free to browse, download, and use at no cost for scientific and educational purposes). The dataset is offered to the radiomics community to compare simple features extracted with different software pipelines as well as to identify features that may not be stable with respect to image acquisition conditions even under highly simplified conditions. Our unique contribution to the field is to investigate the robustness of each radiomic feature with respect to different scanning acquisition parameters.

### ACKNOWLEDGMENTS

The authors thank the in-kind contribution of the commercial vendors, Computerized Imaging Reference Systems (CIRS) and The Phantom Laboratory, that supported our study with the loan of the above mentioned phantoms. The commercial partners have made no direct contribution to the writing of this article. This work has been carried out as part of the Dutch STW-Perspectief Research Program (STRaTeGy grant numbers 14929 and 14930). The national research infrastructure TraIT is being financially supported by the Dutch Cancer Society.

### CONFLICT OF INTEREST

The authors declare no conflict of interests pertaining to the above scientific work.



<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: petros.kalendralis@maastro.nl.

## REFERENCES

1. Fass L. Imaging and cancer: a review. *Mol Oncol.* 2008;2:115–152.
2. Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJR. Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. *Eur J Nucl Med Mol Imaging.* 2013;40:133–140.
3. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48:441–446.
4. Das AK, Bell MH, Nirodi CS, Story MD, Minna JD. Radiogenomics predicting tumor responses to radiotherapy in lung cancer. *Semin Radiat Oncol.* 2010;20:149–155.
5. Larue RTHM, Defraene G, De Ruyscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br J Radiol.* 2017;90:20160665.
6. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magn Reson Imaging.* 2012;30:1234–1248.
7. Miles KA. How to use CT texture analysis for prognostication of non-small cell lung cancer. *Cancer Imaging.* 2016;16:10.
8. Yip SSF, Aerts HJWL. Applications and limitations of radiomics. *Phys Med Biol.* 2016;61:R150–R166.
9. Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. *J Med Imaging.* 2015;2:041002.
10. Fave X, Mackin D, Yang J, et al. Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer? Can radiomics features be measured from CBCT images? *Med Phys.* 2015;42:6784–6797.
11. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* 2015;350:g7594–g7594.
12. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and reproducibility of radiomic features: a systematic review. *Int J Radiat Oncol.* 2018;102:1143–1158.
13. Zwanenburg A, Lööck S. Why validation of prognostic models matters? *Radiother Oncol.* 2018;127:370–373.
14. Hu P, Wang J, Zhong H, et al. Reproducibility with repeat CT in radiomics study for rectal cancer. *Oncotarget.* 2016;7:71440–71446.
15. Balagurunathan Y, et al. Reproducibility and prognosis of quantitative features extracted from CT images. *Transl Oncol.* 2014;7:72–87.
16. Midya A, Chakraborty J, Gönen M, Do RKG, Simpson AL. Influence of CT acquisition and reconstruction parameters on radiomic feature reproducibility. *J Med Imaging.* 2018;5:1.
17. Zhao B, Tan Y, Tsai WY, Schwartz LH, Lu L. Exploring variability in CT characterization of tumors: a preliminary phantom study. *Transl Oncol.* 2014;7:88–93.
18. Catphan® 700 Manual. The Phantom Laboratory; 2016.
19. Catphan® 700 Data Sheet. The Phantom Laboratory; 2016.
20. CTP698 and CCT162 COPDGene Lung Phantom II Data Sheet. The Phantom Laboratory; 2014.
21. CTP698 and CCT162 Lung Phantom II Manual. The Phantom Laboratory; 2014.
22. Triple modality 3D Abdominal Phantom Data Sheet. Computerized Imaging Reference Systems (CIRS); 2013.
23. Marcus DS, Olsen TR, Ramaratnam M, Buckner RL. The extensible neuroimaging archive toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. *Neuroinformatics.* 2007;5:11–34.
24. Archie KA, Marcus DS. DicomBrowser: software for viewing and modifying DICOM metadata. *J Digit Imaging.* 2012;25:635–645.
25. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017;77:e104–e107.
26. Zwanenburg A, Leger S, Vallières M, Lööck S, for the I. B. S. Initiative. Image biomarker standardisation initiative; 2016. *ArXiv161207003 Cs.*
27. Johan VS, Lustberg T, Grittner D, et al. Towards a semantic PACS: using semantic web technology to represent imaging data. *Stud Health Technol Inform.* 2014;205:166–170.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Data S1:** Supplementary material with all the information used for the analysis of the scans of the quality assurance phantom Catphan 700.